Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
L1	0	inlink$3 and outlink$3 and (content adj based or content near3 based) and (relevance near3 rank)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2005/09/12 17:02
L2	77	(content adj based or content near3 based) and (relevance near3 rank)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2005/09/12 17:03
L3	1	((content adj based or content near3 based) and (relevance near3 rank)).ti.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2005/09/12 17:02
L4	1	((content adj based or content near3 based) and (relevance near3 rank)).ab.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2005/09/12 17:02
L5	2	((content adj based or content near3 based) and (relevance near3 rank)).clm.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2005/09/12 17:02
L6	0	((content adj based or content near3 based) and (relevance near3 rank)) and (inlink$3 near3 outlink$3)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2005/09/12 17:04
L7	64	"707"/$.ccls. and ((content adj based or content near3 based) and (relevance near3 rank))	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2005/09/12 17:11
L8	25	"707"/$.ccls. and (inlink$3 near3 outlink$3)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2005/09/12 17:05
L9	3	("707"/$.ccls. and (inlink$3 near3 outlink$3)) and (probability near3 value) and (web same page)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2005/09/12 17:06

L10	0	(709/218,219.ccls. and (inlink$3 near3 outlink$3)) and (probability near3 value) and (web same page)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2005/09/12 17:12
L11	4474	709/218,219.ccls.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2005/09/12 17:07
L12	14	11 and (probability near3 value) and (web same page)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2005/09/12 17:08
L13	0	8 and 12	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2005/09/12 17:08
L14	2424	"content-based" or (content near based) and "relevance rank" and "web page"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2005/09/12 17:10
L15	21	7 and 14	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2005/09/12 17:11
L16	1	(709/203,219.ccls. and (inlink$3 near3 outlink$3)) and (probability near3 value) and (web same page)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2005/09/12 17:12

PORTAL USPTO

Subscribe (Full Service) Register (Limited Service, Free) Login

Search: The ACM Digital Library The Guide

"inlink" +"outlink" "web page" "probility value"

THE ACM DIGITAL LIBRARY

Feedback Report a problem Satisfaction survey

Terms used **inlink outlink web page probility value** Found **83** of **160,906**

Sort results by: relevance

Display results: expanded form

Save results to a Binder

Search Tips

Open results in a new window

Try an Advanced Search

Try this search in The ACM Guide

Results 1 - 20 of 83 Result page: **1** 2 3 4 5 next

Relevance scale

1 Emergent web patterns: The connectivity sonar: detecting site functionality by structural patterns

Einat Amitay, David Carmel, Adam Darlow, Ronny Lempel, Aya Soffer

August 2003 **Proceedings of the fourteenth ACM conference on Hypertext and hypermedia**

Full text available: pdf(153.40 KB) Additional Information: full citation, abstract, references, citings, index terms

Web sites today serve many different functions, such as corporate sites, search engines, e-stores, and so forth. As sites are created for different purposes, their structure and connectivity characteristics vary. However, this research argues that sites of similar role exhibit similar structural patterns, as the functionality of a site naturally induces a typical hyperlinked structure and typical connectivity patterns to and from the rest of the Web. Thus, the functionality of Web sites is refle ...

Keywords: link analysis, web IR, web graphs

2 Link Analysis: Improvement of HITS-based algorithms on web documents

Longzhuang Li, Yi Shang, Wei Zhang

May 2002 **Proceedings of the 11th international conference on World Wide Web**

Full text available: pdf(214.35 KB) Additional Information: full citation, abstract, references, citings, index terms

In this paper, we present two ways to improve the precision of HITS-based algorithms on Web documents. First, by analyzing the limitations of current HITS-based algorithms, we propose a new weighted HITS-based method that assigns appropriate weights to in-links of root documents. Then, we combine content analysis with HITS-based algorithms and study the effects of four representative relevance scoring methods, **VSM, Okapi, TLS**, and **CDR**, using a set of broad topic queries. Our experi ...

Keywords: HITS-based algorithms, information retrieval, relevance scoring methods

3 Re-ranking search results using network analysis a case study with google: a case study with Google

Behnak Yaltaghian, Mark Chignell

September 2002 **Proceedings of the 2002 conference of the Centre for Advanced**

Studies on Collaborative research

Full text available: pdf(186.14 KB) Additional Information: full citation, abstract, references, citings, index terms

In this paper we review methods of structured search for information on the World Wide Web. We propose new methods based on co-citation and network analysis. We describe a set of 21 measures based on these methods and examine the factor structure of those measures. We then report on a recent study that we have conducted at the University of Toronto. Human judges rated the relevance of a selection of Web pages returned by the Google search engine for each of seven queries. We compared the average ...

4 Intelligent crawling on the World Wide Web with arbitrary predicates

Charu C. Aggarwal, Fatima Al-Garawi, Philip S. Yu

April 2001 **Proceedings of the 10th international conference on World Wide Web**

Full text available: pdf(272.60 KB) Additional Information: full citation, references, citings, index terms

Keywords: World Wide Web, crawling, querying

5 Link-based ranking 2: Adaptive ranking of web pages

Ah Chung Tsoi, Gianni Morini, Franco Scarselli, Markus Hagenbuchner, Marco Maggini

May 2003 **Proceedings of the 12th international conference on World Wide Web**

Full text available: pdf(1.48 MB) Additional Information: full citation, abstract, references, citings, index terms

In this paper, we consider the possibility of altering the PageRank of web pages, from an administrator's point of view, through the modification of the PageRank equation. It is shown that this problem can be solved using the traditional quadratic programming techniques. In addition, it is shown that the number of parameters can be reduced by clustering web pages together through simple clustering techniques. This problem can be formulated and solved using quadratic programming techniques. It is ...

Keywords: PageRank, adaptive PageRank determinations, learning PageRank, quadratic programming applications, search engine

6 Web Information Retrieval: The Importance of Prior Probabilities for Entry Page Search

Wessel Kraaij, Thijs Westerveld, Djoerd Hiemstra

August 2002 **Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval**

Full text available: pdf(135.87 KB) Additional Information: full citation, abstract, references, citings, index terms

An important class of searches on the world-wide-web has the goal to find an entry page (homepage) of an organisation. Entry page search is quite different from Ad Hoc search. Indeed a plain Ad Hoc system performs disappointingly. We explored three non-content features of web pages: page length, number of incoming links and URL form. Especially the URL form proved to be a good predictor. Using URL form priors we found over 70% of all entry pages at rank 1, and up to 89% in the top 10. Non-conten ...

Keywords: URLs, entry page search, language models, links, parameter estimation, prior probabilities

7 Industrial and practical experience track paper session 2: Crawling a country: better

strategies than breadth-first for web page ordering
Ricardo Baeza-Yates, Carlos Castillo, Mauricio Marin, Andrea Rodriguez
May 2005 **Special interest tracks and posters of the 14th international conference on World Wide Web**
Full text available: pdf(275.52 KB) Additional Information: full citation, abstract, references, index terms

This article compares several page ordering strategies for Web crawling under several metrics. The objective of these strategies is to download the most "important" pages "early" during the crawl. As the coverage of modern search engines is small compared to the size of the Web, and it is impossible to index all of the Web for both theoretical and practical reasons, it is relevant to index at least the most important pages.We use data from actual Web pages to build Web graphs and execute a crawl ...

Keywords: scheduling policy, web crawler, web page importance

8 Web Behavior Patterns: Separating the swarm: categorization methods for user sessions on the web
Jeffrey Heer, Ed H. Chi
April 2002 **Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves**
Full text available: pdf(462.60 KB) Additional Information: full citation, abstract, references, citings, index terms

Understanding user behaviors on Web sites enables site owners to make sites more usable, ultimately helping users to achieve their goals more quickly. Accordingly, researchers have devised methods for categorizing user sessions in hopes of revealing user interests. These techniques build user profiles by combining users' navigation paths with other data features, such as page viewing time, hyperlink structure, and page content. Previously, we have presented complex techniques of combining many o ...

Keywords: World Wide Web, classification, clustering, data mining, user categorization, user patterns, user profile, user study, web mining

9 Silk from a sow's ear: extracting usable structures from the Web
Peter Pirolli, James Pitkow, Ramana Rao
April 1996 **Proceedings of the SIGCHI conference on Human factors in computing systems: common ground**
Full text available: pdf(1.26 MB) html(35.72 KB) Additional Information: full citation, references, citings, index terms

Keywords: World Wide Web, hypertext, information visualization

10 Link analysis: Ranking the web frontier
Nadav Eiron, Kevin S. McCurley, John A. Tomlin
May 2004 **Proceedings of the 13th international conference on World Wide Web**
Full text available: pdf(238.97 KB) Additional Information: full citation, abstract, references, index terms

The celebrated PageRank algorithm has proved to be a very effective paradigm for ranking results of web search algorithms. In this paper we refine this basic paradigm to take into account several evolving prominent features of the web, and propose several algorithmic innovations. First, we analyze features of the rapidly growing "frontier" of the web, namely the part of the web that crawlers are unable to cover for one reason or another. We analyze

the effect of these pages and find it to be sig ...

Keywords: hypertext, pagerank, ranking

11 Effect of different network analysis strategies on search engine re-ranking
Behnak Yaltaghian, Mark H. Chignell
October 2004 **Proceedings of the 2004 conference of the Centre for Advanced Studies on Collaborative research**
Full text available: pdf(253.28 KB) Additional Information: full citation, abstract, references, index terms

The research described in this paper examined two different approaches to building the co-citation network that the authors have used in re-ranking the set of results returned by a search engine [22, 23]. The more computationally demanding (in terms of query load) Inter- or Web-wide co-citation approach used in-links from throughout the Web to build the network. In contrast, the Intra co-citation approach only used inlinks inferred from search engine output. Results of this study confirmed th ...

12 Papers: Do TREC web collections look like the web?
Ian Soboroff
September 2002 **ACM SIGIR Forum**, Volume 36 Issue 2
Full text available: pdf(289.72 KB) Additional Information: full citation, abstract, references

We measure the WT10g test collection, used in the TREC-9 and TREC 2001 Web Tracks, and the .GOV test collection used in the TREC 2002 Web and Interactive Tracks, with common measures used in the web topology community, in order to see if these collections "look like" the web. This is not an idle question; characteristics of the web, such as power law relationships, diameter, and connected components have all been observed within the scope of general web crawls, constructed by blindly following l ...

13 Web search 1: Exploiting the hierarchical structure for link analysis
Gui-Rong Xue, Qiang Yang, Hua-Jun Zeng, Yong Yu, Zheng Chen
August 2005 **Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval SIGIR '05**
Full text available: pdf(562.55 KB) Additional Information: full citation, abstract, references, index terms

Link analysis algorithms have been extensively used in Web information retrieval. However, current link analysis algorithms generally work on a flat link graph, ignoring the hierarchal structure of the Web graph. They often suffer from two problems: the sparsity of link graph and biased ranking of newly-emerging pages. In this paper, we propose a novel ranking algorithm called Hierarchical Rank as a solution to these two problems, which considers both the hierarchical structure and the link stru ...

Keywords: hierarchical random walk model, hierarchical web graph, link analysis

14 Information retrieval session 7: web: Combining link-based and content-based methods for web document classification
Pável Calado, Marco Cristo, Edleno Moura, Nivio Ziviani, Berthier Ribeiro-Neto, Marcos André Gonçalves
November 2003 **Proceedings of the twelfth international conference on Information and knowledge management**
Full text available: pdf(206.14 KB) Additional Information: full citation, abstract, references, citings, index terms

This paper studies how link information can be used to improve classification results for Web collections. We evaluate four different measures of subject similarity, derived from the

Web link structure, and determine how accurate they are in predicting document categories. Using a Bayesian network model, we combine these measures with the results obtained by traditional content-based classifiers. Experiments on a Web directory show that best results are achieved when links from pages outside the ...

Keywords: Bayesian networks, classification, link analysis, web

15 Link analysis: Sic transit gloria telae: towards an understanding of the web's decay
Ziv Bar-Yossef, Andrei Z. Broder, Ravi Kumar, Andrew Tomkins
May 2004 **Proceedings of the 13th international conference on World Wide Web**

Full text available: pdf(248.69 KB) Additional Information: full citation, abstract, references, index terms

The rapid growth of the web has been noted and tracked extensively. Recent studies have however documented the dual phenomenon: web pages have small half lives, and thus the web exhibits rapid death as well. Consequently, page creators are faced with an increasingly burdensome task of keeping links up-to-date, and many are falling behind. In addition to just individual pages, collections of pages or even entire neighborhoods of the web exhibit significant *decay*, rendering them less effect ...

Keywords: 404 return code, dead links, link analysis, web decay, web information retrieval

16 Posters: Outlink estimation for pagerank computation under missing data
Sreangsu Acharyya, Joydeep Ghosh
May 2004 **Proceedings of the 13th international World Wide Web conference on Alternate track papers & posters**

Full text available: pdf(79.17 KB) Additional Information: full citation, abstract, references, index terms

The enormity and rapid growth of the web-graph forces quantities such as its pagerank tobe computed under missing information consisting of outlinks of pages that have not yet been crawled. This paper examines the role played by the size and distribution of this missing data in determining the accuracy of the computed pagerank, focusing on questions such as (i) the accuracy of pageranks under missing information, (ii) the size at which a crawl process may be aborted while still ensuring reasonab ...

17 Industrial and practical experience track paper session 1: Identifying link farm spam pages
Baoning Wu, Brian D. Davison
May 2005 **Special interest tracks and posters of the 14th international conference on World Wide Web**

Full text available: pdf(260.52 KB) Additional Information: full citation, abstract, references, index terms

With the increasing importance of search in guiding today's web traffic, more and more effort has been spent to create search engine spam. Since link analysis is one of the most important factors in current commercial search engines' ranking systems, new kinds of spam aiming at links have appeared. Building link farms is one technique that can deteriorate link-based ranking algorithms. In this paper, we present algorithms for detecting these link farms automatically by first generating a seed se ...

Keywords: HITS, PageRank, link analysis, spam, web search engine

18 PageCluster: Mining conceptual link hierarchies from Web log files for adaptive Web site navigation

Jianhan Zhu, Jun Hong, John G. Hughes
May 2004 **ACM Transactions on Internet Technology (TOIT)**, Volume 4 Issue 2

Full text available: pdf(280.84 KB) Additional Information: full citation, abstract, references, citings, index terms

User traversals on hyperlinks between Web pages can reveal semantic relationships between these pages. We use user traversals on hyperlinks as weights to measure semantic relationships between Web pages. On the basis of these weights, we propose a novel method to put Web pages on a Web site onto different conceptual levels in a link hierarchy. We develop a clustering algorithm called PageCluster, which clusters conceptually-related pages on each conceptual level of the link hierarchy based on th ...

Keywords: Link hierarchies, Web site navigation, bibliographic analysis, clustering, conceptual link hierarchies, link similarity

19 Search engineering 1: Impact of search engines on page popularity
Junghoo Cho, Sourashis Roy
May 2004 **Proceedings of the 13th international conference on World Wide Web**

Full text available: pdf(172.22 KB) Additional Information: full citation, abstract, references, citings, index terms

Recent studies show that a majority of Web page accesses are referred by search engines. In this paper we study the widespread use of Web search engines and its impact on the ecology of the Web. In particular, we study how much impact search engines have on the popularity evolution of Web pages. For example, given that search engines return currently popular" pages at the top of search results, are we somehow penalizing newly created pages that are not very well known yet? Are popular pages gett ...

Keywords: change in pagerank, pagerank, random surfer model, search engine's impact, web evolution

20 Semantic querying: Algorithmic detection of semantic similarity
Ana G. Maguitman, Filippo Menczer, Heather Roinestad, Alessandro Vespignani
May 2005 **Proceedings of the 14th international conference on World Wide Web**

Full text available: pdf(4.10 MB) Additional Information: full citation, abstract, references, index terms

Automatic extraction of semantic information from text and links in Web pages is key to improving the quality of search results. However, the assessment of automatic semantic measures is limited by the coverage of user studies, which do not scale with the size, heterogeneity, and growth of the Web. Here we propose to leverage human-generated metadata --- namely topical directories --- to measure semantic relationships among massive numbers of pairs of Web pages or topics. The Open Directory Proj ...

Keywords: Web mining, Web search, content and link similarity, ranking evaluation, semantic similarity

Results 1 - 20 of 83 Result page: **1** 2 3 4 5 next

The ACM Portal is published by the Association for Computing Machinery. Copyright © 2005 ACM, Inc.
Terms of Usage Privacy Policy Code of Ethics Contact Us

Useful downloads: Adobe Acrobat QuickTime Windows Media Player Real Player